

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



FILE

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/52, 9/00, 15/82, 5/10 // A01H 5/00		(11) International Publication Number: WO 99/53072
A1		(43) International Publication Date: 21 October 1999 (21.10.99)
(21) International Application Number: PCT/US99/07639		(81) Designated States: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 8 April 1999 (08.04.99)		
(30) Priority Data: 60/081,143 9 April 1998 (09.04.98) US		
(71) Applicant (for all designated States except US): E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).		
(72) Inventors; and (75) Inventors/Applicants (for US only): CRESSMAN, Robert, F. [US/US]; 2518 Dean Drive, Wilmington, DE 19808 (US). ALLEN, Stephen, M. [US/US]; 2225 Rosewood Drive, Wilmington, DE 19810 (US).		
(74) Agent: MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).		

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

19 10 21 10

(54) Title: STARCH R1 PHOSPHORYLATION PROTEIN HOMOLOGS

```
1 60
SEQ ID NO:15 (gi3287270) MSNSLGNLLYQGFLTSTVLEHKSRLSPPCVGGNSLFQQQVISKSPLEFRGNRLKVQK
SEQ ID NO:10 -----
SEQ ID NO:12 -----

61 120
SEQ ID NO:15 (gi3287270) KKIPMEKKRAFSSSPHAVLTTDTSSSLAEKPSLGGNIELQVDVRPPTSGDVSFVDFQVTN
SEQ ID NO:10 -----
SEQ ID NO:12 -----

121 180
SEQ ID NO:15 (gi3287270) GSDKLFRLWCAVKFKETNSLPNDRPDGTVKYKKAALRTPFVKSGSNSILRLRIDTALE
SEQ ID NO:10 -----
SEQ ID NO:12 -----

181 240
SEQ ID NO:15 (gi3287270) AIEFLIYDEANDKWKNNNGGNFRVKLSRKEIRGPOVSVPEELVQIQSYLRWERKGRQWYP
SEQ ID NO:10 -----
SEQ ID NO:12 -----

241 300
SEQ ID NO:15 (gi3287270) PEKEKEEYFAARTVLQEEIARGASIQDIRARLTNTNOKSQSKEEPLHVTKSDIPDDLAQA
SEQ ID NO:10 -----
SEQ ID NO:12 -----

301 360
SEQ ID NO:15 (gi3287270) QAYIRWEKAGKPNYPPEKQIEELEPARRELQLELKGITLDELKRTITKGEIKTKVERHL
SEQ ID NO:10 -----
SEQ ID NO:12 -----

361 420
SEQ ID NO:15 (gi3287270) KRSSFVERIQRKKRDFGHLINKYTSSPAVQVQKVL EEPALSKIKLYAKEKEEQIDQPI
SEQ ID NO:10 -----
SEQ ID NO:12 -----
```

(57) Abstract

This invention relates to an isolated nucleic acid fragment encoding a starch R1 phosphorylation protein. The invention also relates to the construction of a chimeric gene encoding all or a portion of the starch R1 phosphorylation protein in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the starch R1 phosphorylation protein in a transformed host cell.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE

STARCH R1 PHOSPHORYLATION PROTEIN HOMOLOGS

This application claims the benefit of U.S. Provisional Application No. 60/081,143, filed April 9, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding starch R1 phosphorylation proteins in plants and seeds.

BACKGROUND OF THE INVENTION

Starch is a mixture of two polysaccharides, amylose and amylopectin. Amylose is an unbranched chain of up to several thousand α -D-glucopyranose units linked by α -1,4 glycosidic bonds. Amylopectin is a highly branched molecule made up of up to 50,000 α -D-glucopyranose residues linked by α -1,4 and α -1,6 glycosidic bonds. Approximately 5% of the glycosidic linkages in amylopectin are α -1,6 bonds, which leads to the branched structure of the polymer.

Amylose and amylopectin molecules are organized into granules that are stored in plastids. The starch granules produced by most plants are 15-30% amylose and 70-85% amylopectin. The ratio of amylose to amylopectin and the degree of branching of amylopectin affects the physical and functional properties of the starch. Functional properties, such as viscosity and stability of a gelatinized starch determine the usefulness and hence the value of starches in food and industrial applications.

The R1 protein of potato appears to be a granule associated enzyme that is involved in starch phosphorylation (Lorberth, R. et al. (1998) *Nature Biotechnology* 16:473-477). R1 activity has been associated with starch degradation in potato tubers. Studies have shown that inhibition of R1 activity leads to the synthesis of modified starch that is not degraded by enzymes present in potato tissue. If changes in starch degradation are a direct consequence of changes in the degree of phosphorylation this suggests that starch phosphorylation is an important modification that promotes degradation.

Accordingly, the availability of nucleic acid sequences encoding all or a portion of R1 proteins in other plants would facilitate studies to better understand starch degradation and provide genetic tools for the manipulation of starch biosynthesis plant cells.

SUMMARY OF THE INVENTION

The instant invention relates to isolated nucleic acid fragments encoding starch R1 phosphorylation proteins. Specifically, this invention concerns isolated nucleic acid fragments encoding starch R1 phosphorylation proteins from *Arabidopsis*, ginger, moss, cattail, rice and soybean. In addition, this invention relates to a nucleic acid fragment that is complementary to the nucleic acid fragments encoding starch R1 phosphorylation proteins.

An additional embodiment of the instant invention pertains to a polypeptide encoding all or a substantial portion of a starch R1 phosphorylation protein.

In another embodiment, the instant invention relates to a chimeric gene encoding a starch R1 phosphorylation protein, or to a chimeric gene that comprises a nucleic acid fragment that is complementary to a nucleic acid fragment encoding a starch R1 phosphorylation protein, operably linked to suitable regulatory sequences, wherein expression of the chimeric gene results in production of levels of the encoded protein in a transformed host cell that is altered (i.e., increased or decreased) from the level produced in an untransformed host cell.

In a further embodiment, the instant invention concerns a transformed host cell comprising in its genome a chimeric gene encoding a starch R1 phosphorylation protein operably linked to suitable regulatory sequences. Expression of the chimeric gene results in production of altered levels of the encoded protein in the transformed host cell. The transformed host cell can be of eukaryotic or prokaryotic origin, and include cells derived from higher plants and microorganisms. The invention also includes transformed plants that arise from transformed host cells of higher plants, and seeds derived from such transformed plants.

An additional embodiment of the instant invention concerns a method of altering the level of expression of a starch R1 phosphorylation protein in a transformed host cell comprising: a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a starch R1 phosphorylation protein; and b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of altered levels of starch R1 phosphorylation protein in the transformed host cell.

An additional embodiment of the instant invention concerns a method for obtaining a nucleic acid fragment encoding all or a substantial portion of an amino acid sequence encoding a starch R1 phosphorylation protein.

BRIEF DESCRIPTION OF THE DRAWING AND SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing which form a part of this application.

Figure 1 shows a comparison of the amino acid sequences of potato starch R1 phosphorylation protein set forth in NCBI Identifier No. gi 3287270 (SEQ ID NO:15), the instant rice R1 homolog (SEQ ID NO:10) and the instant soybean R1 homolog (SEQ ID NO:12).

The following sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

SEQ ID NO:1 is the nucleotide sequence comprising a portion of the cDNA insert in clone acs2c.pk001.g20 encoding a portion of an *Arabidopsis* starch R1 phosphorylation protein.

5 SEQ ID NO:2 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:1.

SEQ ID NO:3 is the nucleotide sequence comprising a portion of the cDNA insert in clone ecr1c.pk007.119 encoding a portion of a ginger (*Curcuma zedoaria*) starch R1 phosphorylation protein.

10 SEQ ID NO:4 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:3.

SEQ ID NO:5 is the nucleotide sequence comprising a portion of the cDNA insert in clone emm1c.pk001.p18 encoding a portion of a moss (*Brachythecium oxycladon*, *Plagiomnium cuspidatum* and *Amblystegium varium*) starch R1 phosphorylation protein.

15 SEQ ID NO:6 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:5.

SEQ ID NO:7 is the nucleotide sequence comprising a portion of the cDNA insert in clone etr1c.pk003.c21 encoding a portion of a cattail (*Typha latifolia*) starch R1 phosphorylation protein.

20 SEQ ID NO:8 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:7.

SEQ ID NO:9 is the nucleotide sequence comprising a contig assembled from the cDNA inserts in clones rlm4n.pk003.p17, rl0n.pk088.j11 and rlr6.pk0099.d9 encoding a portion of a rice starch R1 phosphorylation protein.

25 SEQ ID NO:10 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:9.

SEQ ID NO:11 is the nucleotide sequence comprising a contig assembled from the cDNA inserts in clones scr1c.pk003.e3, ses4d.pk0019.b5, sl1.pk0109.f9, sl2.pk0041.d7, src3c.pk006.d11 and src3c.pk026.j6 encoding a portion of a soybean starch R1 phosphorylation protein.

30 SEQ ID NO:12 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:11.

SEQ ID NO:13 is the nucleotide sequence comprising a portion of the cDNA insert in clone scr1c.pk002.k14 encoding a portion of a soybean starch R1 phosphorylation protein.

35 SEQ ID NO:14 is the deduced amino acid sequence of a portion of a starch R1 phosphorylation protein derived from the nucleotide sequence of SEQ ID NO:13.

SEQ ID NO:15 is the amino acid sequence of potato tuber starch R1 phosphorylation protein, (NCBI Identifier No. gi 3287270).

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Research* 13:3021-3030 (1985) and in the *Biochemical Journal* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, an "isolated nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA. As used herein, "contig" refers to an assemblage of overlapping nucleic acid sequences to form one contiguous nucleotide sequence. For example, several DNA sequences can be compared and aligned to identify common or overlapping regions. The individual sequences can then be assembled into a single contiguous nucleotide sequence.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the protein encoded by the DNA sequence.

"Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate alteration of gene expression by antisense or co-suppression technology or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary sequences.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a gene which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded protein, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one

positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize, under stringent conditions (0.1X SSC, 0.1% SDS, 65°C), with the nucleic acid fragments disclosed herein.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent similarity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Preferred are those nucleic acid fragments whose nucleotide sequences encode amino acid sequences that are 80% similar to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% similar to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% similar to the amino acid sequences reported herein. Sequence alignments and percent similarity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins, D. G. and Sharp, P. M. (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10), (hereafter Clustal algorithm). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to afford putative identification of that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as

amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches partial or complete amino acid and nucleotide sequences encoding one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment that encodes all or a substantial portion of the amino acid sequence encoding the starch R1 phosphorylation proteins as set forth in SEQ ID NOs:2, 4, 6, 8, 10, 12 and 14. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic genes" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments which are then enzymatically assembled to construct the entire gene. "Chemically synthesized", as related to a sequence of DNA, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of DNA may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the genes can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived

from the same source, but arranged in a manner different than that found in nature.

"Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

"Coding sequence" refers to a DNA sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

"Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg, (1989) *Biochemistry of Plants* 15:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a DNA sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D. (1995) *Molecular Biotechnology* 3:225).

The "3' non-coding sequences" refer to DNA sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The

polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al., (1989) *Plant Cell* 1:671-680.

5 "RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a
10 double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to RNA transcript that includes the mRNA and so can be translated into protein by the cell. "Antisense RNA" refers to a RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an
15 antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

20 The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

25 The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production
30 of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

35 "Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and

propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

5 A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels, J. J., (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*)
10 can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included (Raikhel (1992) *Plant Phys.* 100:1627-1632).

"Transformation" refers to the transfer of a nucleic acid fragment into the genome of a
15 host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol.* 143:277) and particle-accelerated or "gene gun" transformation technology (Klein et al. (1987) *Nature (London)* 327:70-73; U.S. Patent No. 4,945,050,
20 incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Maniatis").

25 Nucleic acid fragments encoding at least a portion of several starch R1 phosphorylation proteins have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. Table 1 lists the proteins that are described herein, and the designation of the cDNA clones that comprise the nucleic acid
30 fragments encoding these proteins.

TABLE 1

Starch R1 Phosphorylation Proteins

Enzyme	Clone	Plant
Starch R1 Phosphorylation Protein	acs2c.pk001.g20	<i>Arabidopsis</i>
	ecr1c.pk007.l19	Ginger
	emmlc.pk001.p18	Moss
	etr1c.pk003.c21	Cattail
	rlm4n.pk003.p17	Rice
	rl0n.pk088.j11	Rice
	rlr6.pk0099.d9	Rice
	scr1c.pk003.e3	Soybean
	ses4d.pk0019.b5	Soybean
	sl1.pk0109.f9	Soybean
	sl2.pk0041.d7	Soybean
	src3c.pk006.d11	Soybean
	src3c.pk026.j6	Soybean
	scr1c.pk002.k14	Soybean

5 The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain
10 reaction).

For example, genes encoding other starch R1 phosphorylation proteins, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific
15 oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed
20 and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al., (1988) *PNAS USA* 85:8998) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al., (1989) *PNAS USA* 86:5673; Loh et al., (1989) *Science* 243:217). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman, M. A. and Martin, G. R., (1989) *Techniques* 1:165).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner, R. A. (1984) *Adv. Immunol.* 36:1; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed starch R1 phosphorylation proteins are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of starch phosphorylation in those cells.

Overexpression of the starch R1 phosphorylation proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then be constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host

plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al., (1985) *EMBO J.* 4:2411-2418; De Almeida et al., (1989) *Mol. Gen. Genetics* 218:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant starch R1 phosphorylation proteins to different cellular compartments, or to facilitate its secretion from the cell. It is thus envisioned that the chimeric gene described above may be further supplemented by altering the coding sequence to encode a starch R1 phosphorylation protein with appropriate intracellular targeting sequences such as transit sequences (Keegstra, K. (1989) *Cell* 56:247-253), signal sequences or sequences encoding endoplasmic reticulum localization (Chrispeels, J. J., (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53), or nuclear localization signals (Raikhel, N. (1992) *Plant Phys.* 100:1627-1632) added and/or with targeting sequences that are already present removed. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of utility may be discovered in the future.

It may also be desirable to reduce or eliminate expression of genes encoding starch R1 phosphorylation proteins in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant starch R1 phosphorylation proteins can be constructed by linking a gene or gene fragment encoding an starch R1 phosphorylation protein to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

The instant starch R1 phosphorylation proteins (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting starch R1 phosphorylation proteins *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant starch R1 phosphorylation proteins are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to

construct a chimeric gene for production of the instant starch R1 phosphorylation proteins. This chimeric gene could then be introduced into appropriate microorganisms via transformation to provide high level expression of the encoded starch R1 phosphorylation protein. An example of a vector for high level expression of the instant starch R1 phosphorylation proteins in a bacterial host is provided (Example 6).

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al., (1987) *Genomics* 1:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein, D. et al., (1980) *Am. J. Hum. Genet.* 32:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in R. Bernatzky, R. and Tanksley, S. D. (1986) *Plant Mol. Biol. Reporter* 4(1):37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel, J. D., et al., In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask, B. J. (1991) *Trends Genet.* 7:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan, M. et al. (1995) *Genome Research* 5:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian, H. H. (1989) *J. Lab. Clin. Med.* 114(2):95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield, V. C. et al. (1993) *Genomics* 16:325-332), allele-specific ligation (Landegren, U. et al. (1988) *Science* 241:1077-1080), nucleotide extension reactions (Sokolov, B. P. (1990) *Nucleic Acid Res.* 18:3671), Radiation Hybrid Mapping (Walter, M. A. et al. (1997) *Nature Genetics* 7:22-28) and Happy Mapping (Dear, P. H. and Cook, P. R. (1989) *Nucleic Acid Res.* 17:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer, (1989) *Proc. Natl. Acad. Sci USA* 86:9402; Koes et al., (1995) *Proc. Natl. Acad. Sci USA* 92:8149; Bensen et al., (1995) *Plant Cell* 7:75). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the starch R1 phosphorylation protein. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding a starch R1 phosphorylation protein can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the starch R1 phosphorylation protein gene product.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without

departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

- 5 cDNA libraries representing mRNAs from various Arabidopsis, ginger, moss, rice and soybean tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from *Arabidopsis*, Ginger, Moss, Rice and Soybean

Library	Tissue	Clone
acs2c	<i>Arabidopsis</i> fertilized carpels with developing seeds 6-7 days after fertilization	acs2c.pk001.g20
ecr1c	Ginger (<i>Curcuma zedoaria</i> , aka shoti starch) developing rhizomes	ecr1c.pk007.i19
emmlc	Moss of three variety (<i>Brachythecium oxycladon</i> , <i>Plagiomnium cuspidatum</i> , <i>Amblystegium varium</i>)	emmlc.pk001.p18
etr1c	Cattail (<i>Typha latifolia</i>) root	etr1c.pk003.c21
rlm4n	Rice (<i>Oryza Sativa</i> , YM) leaf mixture*	rlm4n.pk003.p17
rl0n	Rice (<i>Oryza sativa</i> L.) 15 day leaf*	rl0n.pk088.j11
rlr6	Rice (<i>Oryza sativa</i> L.) leaf 15 days after germination, 6 hours after infection of strain <i>Magaporthe grisea</i> 4360-R-62 (AVR2-YAMO); Resistant	rlr6.pk0099.d9
scr1c	Soybean (<i>Glycine max</i> L., 2872) embryogenic suspension culture	scr1c.pk003.e3 scr1c.pk002.k14
ses4d	Soybean (<i>Glycine max</i> L.) embryogenic suspension 4 days after subculture	ses4d.pk0019.b5
sl1	Soybean (<i>Glycine max</i> L.) Two week old developing seedlings treated with water	sl1.pk0109.f9
sl2	Soybean (<i>Glycine max</i> L.) two week old developing seedlings treated with 2.5 ppm chlorimuron	sl2.pk0041.d7
src3c	Soybean (<i>Glycine max</i> L., Bell) 8 day old root inoculated with eggs of cyst nematode <i>Heterodera glycines</i> (Race 14) for 4 days.	src3c.pk006.d11 src3c.pk026.j6

*These libraries were normalized essentially as described in U.S. Patent No. 5,482,845

- 5 cDNA libraries were prepared in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). Conversion of the Uni-ZAP™ XR libraries into plasmid libraries was accomplished according to the protocol provided by Stratagene. Upon conversion, cDNA inserts were contained in the plasmid vector pBluescript. cDNA inserts from randomly picked bacterial colonies containing
- 10 recombinant pBluescript plasmids were amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences or plasmid DNA was prepared from cultured bacterial cells. Amplified insert DNAs or plasmid DNAs were sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams, M. D. et al., (1991) *Science* 252:1651).
- 15 The resulting ESTs were analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

EXAMPLE 2Identification of cDNA Clones

ESTs encoding starch R1 phosphorylation proteins were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish, W. and States, D. J. (1993) *Nature Genetics* 3:266-272 and Altschul, Stephen F., et al. (1997) *Nucleic Acids Res.* 25:3389-3402) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3Characterization of cDNA Clones Encoding Starch R1 phosphorylation Protein

The BLASTX search using the EST sequences from clones acs2c.pk001.g20, ecrlc.pk007.l19, emmlc.pk001.p18, etrlc.pk003.c21, rlm4n.pk003.p17, rl0n.pk088.j11, rlr6.pk0099.d9, scr1c.pk003.e3, ses4d.pk0019.b5, sl1.pk0109.f9, sl2.pk0041.d7, src3c.pk006.d11, src3c.pk026.j6 and scr1c.pk002.k14 revealed similarity of the proteins encoded by the cDNAs to starch R1 phosphorylation protein from *Solanum tuberosum* (NCBI Identifier No. gi 3287270).

In the process of comparing the ESTs it was found that rice clones rlm4n.pk003.p17, rl0n.pk088.j11 and rlr6.pk0099.d9 had overlapping regions of homology. Soybean clones scr1c.pk003.e3, ses4d.pk0019.b5, sl1.pk0109.f9, sl2.pk0041.d7, src3c.pk006.d11 and src3c.pk026.j6 were also found to have overlapping regions of homology. Using this homology it was possible to align the ESTs and assemble two individual contigs encoding unique rice and soybean starch R1 phosphorylation proteins. The BLAST results for each of the contigs and ESTs are shown in Table 3:

TABLE 3

BLAST Results for Clones Encoding Polypeptides Homologous
to *Solanum tuberosum* Starch R1 phosphorylation Protein

Clone	BLAST pLog Score
acs2c.pk001.g20	62.10
ecrlc.pk007.l19	75.00
emmlc.pk001.p18	56.00
etr1c.pk003.c21	53.00
Contig composed of: rlm4n.pk003.p17 rl0n.pk088.j11 rlr6.pk0099.d9	>250.00
Contig composed of: scr1c.pk003.e3 ses4d.pk0019.b5 sl1.pk0109.f9 sl2.pk0041.d7 src3c.pk006.d11 src3c.pk026.j6	>250.00
scr1c.pk002.k14	94.70

5 The sequence of a portion of the cDNA insert from clone acs2c.pk001.g20 is shown in
SEQ ID NO:1; the deduced amino acid sequence of this cDNA, which represents 11% of the
protein (middle region), is shown in SEQ ID NO:2. The sequence of a portion of the cDNA
insert from clone ecrlc.pk007.l19 is shown in SEQ ID NO:3; the deduced amino acid
sequence of this cDNA, which represents 9.7% of the protein (middle region) is shown in
10 SEQ ID NO:4. The sequence of a portion of the cDNA insert from clone emmlc.pk001.p18
is shown in SEQ ID NO:5; the deduced amino acid sequence of this cDNA, which represents
10.7% of the protein (middle region) is shown in SEQ ID NO:6. The sequence of a portion
of the cDNA insert from clone etrlc.pk003.c21 is shown in SEQ ID NO:7; the deduced
amino acid sequence of this cDNA, which represents 7.7% of the protein (middle region), is
15 shown in SEQ ID NO:8.

The sequence of the rice contig composed of clones rlm4n.pk003.p17, rl0n.pk088.j11
and rlr6.pk0099.d9 is shown in SEQ ID NO:9; the deduced amino acid sequence of this
contig, which represents 33% of the protein (C-terminal region) is shown in SEQ ID NO:10.
Figure 1 presents an alignment of the amino acid sequence set forth in SEQ ID NO:10 and
20 the *Solanum tuberosum* sequence. A calculation of the percent similarity of the amino acid
sequence set forth in SEQ ID NO:10 and the *Solanum tuberosum* sequence (using the Clustal
algorithm) revealed that the protein encoded by the contig is 75.1% similar to the *Solanum
tuberosum* starch R1 phosphorylation protein.

The sequence of the soybean contig composed of clones scr1c.pk003.e3, ses4d.pk0019.b5, sl1.pk0109.f9, sl2.pk0041.d7, src3c.pk006.d11 and src3c.pk026.j6 is shown in SEQ ID NO:11; the deduced amino acid sequence of this contig, which represents 40% of the protein (C-terminal region) is shown in SEQ ID NO:12. Figure 1 presents an alignment of the amino acid sequence set forth in SEQ ID NO:12 and the *Solanum tuberosum* sequence. A calculation of the percent similarity of the amino acid sequence set forth in SEQ ID NO:12 and the *Solanum tuberosum* sequence (using the Clustal algorithm) revealed that the protein encoded by the contig is 76.4% similar to the *Solanum tuberosum* starch R1 phosphorylation protein. The degree of similarity between the rice (SEQ ID NO:10) and soybean amino acid sequences (SEQ ID NO:12) was calculated to be 70.3% (using the Clustal algorithm).

The sequence of a portion of the cDNA insert from clone scr1c.pk002.k14 is shown in SEQ ID NO:13; the deduced amino acid sequence of this cDNA, which represents 11% of the protein (middle region), is shown in SEQ ID NO:14.

BLAST scores and probabilities indicate that the instant nucleic acid fragments encode portions of starch R1 phosphorylation proteins. These sequences represent the first arabidopsis, ginger, moss, cattail, rice and soybean sequences encoding starch R1 phosphorylation proteins.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding a starch R1 phosphorylation protein in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb SalI-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-SalI fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by

restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S.

Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding a starch R1 phosphorylation protein, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al., (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al., (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 µm in diameter) are coated with DNA using the following technique. Ten µg of plasmid DNAs are added to 50 µL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 µL of a 2.5 M solution) and spermidine free base (20 µL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 µL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 µL of ethanol. An aliquot (5 µL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al., (1990) *Bio/Technology* 8:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* 261:9228-9238) can be used for expression of the instant starch R1 phosphorylation protein in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embryos may then be transformed with the expression vector comprising sequences encoding starch R1 phosphorylation proteins. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the

soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Kline et al. (1987) *Nature* (London) 327:70, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al. (1983) *Gene* 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the starch R1 phosphorylation protein and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 µL of a 60 mg/mL 1 µm gold particle suspension is added (in order): 5 µL DNA (1 µg/µL), 20 µL spermidine (0.1 M), and 50 µL CaCl₂ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 µL 70% ethanol and resuspended in 40 µL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five µL of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL

hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant starch R1 phosphorylation proteins can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene* 56:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 µg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 µL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 µg/mL ampicillin. Transformants containing the gene encoding the starch R1 phosphorylation protein are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol.* 189:113-130). Cultures are grown in LB medium

containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio- β -galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μ L of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One μ g of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

CLAIMS

What is claimed is:

1. An isolated nucleic acid fragment encoding all or a substantial portion of a starch R1 phosphorylation protein comprising a member selected from the group consisting of:
 - (a) an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12 and 14;
 - (b) an isolated nucleic acid fragment that is substantially similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12 and 14; and
 - (c) an isolated nucleic acid fragment that is complementary to (a) or (b).
2. The isolated nucleic acid fragment of Claim 1 wherein the nucleotide sequence of the fragment comprises all or a portion of the sequence set forth in a member selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9, 11 and 13.
3. A chimeric gene comprising the nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences.
4. A transformed host cell comprising the chimeric gene of Claim 3.
5. A starch R1 phosphorylation protein polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO: starch R1 phosphorylation protein.
6. A method of altering the level of expression of a starch R1 phosphorylation protein in a host cell comprising:
 - (a) transforming a host cell with the chimeric gene of Claim 3; and
 - (b) growing the transformed host cell produced in step (a) under conditions that are suitable for expression of the chimeric genewherein expression of the chimeric gene results in production of altered levels of a starch R1 phosphorylation protein in the transformed host cell.
7. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a starch R1 phosphorylation protein comprising:
 - (a) probing a cDNA or genomic library with the nucleic acid fragment of Claim 1;
 - (b) identifying a DNA clone that hybridizes with the nucleic acid fragment of Claims 1;
 - (c) isolating the DNA clone identified in step (b); and

- (d) sequencing the cDNA or genomic fragment that comprises the clone isolated in step (c)

wherein the sequenced nucleic acid fragment encodes all or a substantial portion of the amino acid sequence encoding a starch R1 phosphorylation protein.

- 5 8. A method of obtaining a nucleic acid fragment encoding a substantial portion of an amino acid sequence encoding a starch R1 phosphorylation protein comprising:

(a) synthesizing an oligonucleotide primer corresponding to a portion of the sequence set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11 and 13; and

- 10 (b) amplifying a cDNA insert present in a cloning vector using the oligonucleotide primer of step (a) and a primer representing sequences of the cloning vector

wherein the amplified nucleic acid fragment encodes a substantial portion of an amino acid sequence encoding a starch R1 phosphorylation protein

- 15 9. The product of the method of Claim 7.
10. The product of the method of Claim 8.

1 60
SEQ ID NO:15 (gi3287270) MSNSLGNLLYQGFLTSTVLEHKSRISPPCVGNSLFQQQVISKSPSTEFGRNRLKVQK
SEQ ID NO:10
SEQ ID NO:12

61 120
SEQ ID NO:15 (gi3287270) KKIPMEKKRAFSSSPHAVLTDTSSSELAEKFSLGGNIELQVDVRPPTSGDVSFVDFQVTN
SEQ ID NO:10
SEQ ID NO:12

121 180
SEQ ID NO:15 (gi3287270) GSDKLFLHWGAVKFGKETWSLPNDRPDGTVKYKNKALRTPFVKSGSNSILRLEIRDTAIE
SEQ ID NO:10
SEQ ID NO:12

181 240
SEQ ID NO:15 (gi3287270) AIEFLIYDEAHDWKWIKNNGNFRVKLSRKEIRGPDVSVPEELVQIQSYLRWERKQKNYP
SEQ ID NO:10
SEQ ID NO:12

241 300
SEQ ID NO:15 (gi3287270) PEKEKEEYEAARTVLQEEIARGASIQDIRARLTKTNDKSQSKEEPLHVTKSDIPDDLQA
SEQ ID NO:10
SEQ ID NO:12

301 360
SEQ ID NO:15 (gi3287270) QAYIRWEKAGKPNYPPEKQIEELEEARELQLELEKGITLDELKRTITKGEIKTKVEKHL
SEQ ID NO:10
SEQ ID NO:12

361 420
SEQ ID NO:15 (gi3287270) KRSSFaverIQKKRDFGHLINKYTSSPAVQVQVLEPPALSKIKLYAKEKEEQIDDPI
SEQ ID NO:10
SEQ ID NO:12

FIG. 1

421	480
SEQ ID NO:15 (gi3287270)	LNKKFFVDDGELLVLVAKSSGKTKVHLATDLNQPITLHWALSKSPGEMVPPSSILPPG
SEQ ID NO:10	-----
SEQ ID NO:12	-----
481	540
SEQ ID NO:15 (gi3287270)	SIILDKAAETPFSASSDGLTSKVQSLDIVIEDGNFVGMPPFVLLSGEKWIKNQGSDFYVG
SEQ ID NO:10	-----
SEQ ID NO:12	-----
541	600
SEQ ID NO:15 (gi3287270)	FSAASKLALKAAGDGSCTAKSLLDKIAOMESEAKSFMHRFNIAADLIEDATSAGELGFA
SEQ ID NO:10	-----
SEQ ID NO:12	-----
601	660
SEQ ID NO:15 (gi3287270)	GILVMRFMATROLIWNKNYNVKNPREISKAQDRLTDLQNAFTSHPOYREILRMIMSTVG
SEQ ID NO:10	-----
SEQ ID NO:12	-----
661	720
SEQ ID NO:15 (gi3287270)	RGEGDVQIRIDEILVIQRNDCKGGMQEWQKLNNTSPDDVVICQALIDYIKSDFD
SEQ ID NO:10	-----
SEQ ID NO:12	-----
721	780
SEQ ID NO:15 (gi3287270)	LGVIWKTLNENGITKERLLSYDRAIHSEPNFRGDKGGLLRDLGLHYMRTLKAVHSGADLE
SEQ ID NO:10	-----
SEQ ID NO:12	-----
781	840
SEQ ID NO:15 (gi3287270)	SAIANCMGYKTEGEGFMVGVIQNPVSGLPFGQDLLHFVLDHVEDKNVETLLERLLEARE
SEQ ID NO:10	-----
SEQ ID NO:12	-----

FIG. 1 (Continued)

3/4

841 900
SEQ ID NO:15 (gi3287270) ELRPLLLKPNRLKDLLFDIALDSTVRTAVERGYEELNNANPEKIMYFISLVLENLALS
SEQ ID NO:10 -----V-----
SEQ ID NO:12 -----IMYFISLVLENLALS-----

901 960
SEQ ID NO:15 (gi3287270) VDDNEDLVYCLKGWNQALSMSNGDNHWFALFAKAVLDRTRLALASKAEWYHLLQPSAEY
SEQ ID NO:10 -----
SEQ ID NO:12 SDDNEDLIYCLKGWDVALSMCKIKDTHWALYAKSVLDTRLALTNKAHLYQEILQPSAEY-----

961 1020
SEQ ID NO:15 (gi3287270) LGSILGVDQWALNIFTEEIIIRAGSAASLSSLLNRLDPVLRKTANLGSWQIIISPVEAVGYV
SEQ ID NO:10 -----PATLSALLNRIDPVLNRVNAQLGSWQVISPVEVSGYI-----
SEQ ID NO:12 LGSLLGVDKWAVEIFTEEIIIRAGSAASLSTLLNRLDPVLRKTAHLGSWQVISPVEVTVGYV-----

1021 1080
SEQ ID NO:15 (gi3287270) VVVDLLSVQNEIYEKPTILVAKSVKGEETIPDGAVALITPDMPDVLSHVSVRARNGKVC
SEQ ID NO:10 VVVDLLAVQNKSYDKPTILVAKSVKGEETIPDGVGVITPDMPDVLSHVSVRARNCVIL
SEQ ID NO:12 EVIDELLAVQNKSYERPTILIAKSVRGEETIPDGTVAULTPDMPDVLSHVSVRARNSKVC-----

1081 1140
SEQ ID NO:15 (gi3287270) FATCFDPNILLADLOAKEGRILLKPTPSDIIYSEVNEIELQSSS--NLVEAETSATLRLV
SEQ ID NO:10 FATCFDPNTLSELQHDGKVFSEFKPTSAADITYREIPESELQ-SGSLNAEAGQAVPSVSLV
SEQ ID NO:12 FATCFDPNILLANLQENKGLRLKPTSAADVYSEVKEGELIDDKSTQLKDVGSVSPISLA-----

1141 1200
SEQ ID NO:15 (gi3287270) KKQFGGCYAIASADEFTSEMVGAKSRNIAYLKGKVPSSVGIPTSVALPFGVFEKVLSDIN
SEQ ID NO:10 KKKFLGKYAISAEFESEMVGAKSRNVAYLKGKVPSSVGIPTSVAIPFGTFEKVLSDIN
SEQ ID NO:12 RKKFSGRYAVSSEFTGEMVGAKSRNISYLGKVASWIGIPTSVAIPFGVFEHVLSDKPN-----

FIG. 1 (Continued)

1201 1260
SEQ ID NO: 15 (gi3287270) QGVAKELQILMKKLESEGDFSEALGEIRTTVLDSLAPAQVLKELKEKMQSGMPWPGDEGPK
SEQ ID NO: 10 KEVAQTIQMLKGKLAQDDFSALGEIRKTVLNLTAPTQLIKELKEKMLGSGMPWPGDEGQ
SEQ ID NO: 12 QAVAEVNNLKKKLEIEGDFSVLKEIRETVLQNLNAPSHLVEELKTKMKSSGMPWPGDEGEQ
1261 1320
SEQ ID NO: 15 (gi3287270) RWEQAWMAIKKVVASKWNERAYFSTRKVKLDHLYLCMAVLVQEIINADYAFVIHTTNPSS
SEQ ID NO: 10 RWEQAWMAIKKVVAXXWNERAYFSTRKVKLDHLYLSMAVLVQEIIVNADYAFVIHTTNPSS
SEQ ID NO: 12 RWEQAWIAIKKVGWSKWNERAYFSTRKVKLDHLYLSMAVLVQEVINADYAFVIHTTNPAS
1321 1380
SEQ ID NO: 15 (gi3287270) GDDSEIYAEVVRGLGETLVGAYPGRALSFICKKKDLNSPQVLGYPSKPIGLFIKRSIIIFR
SEQ ID NO: 10 GDSSEIYAEVVKGLGETLVGAYPGRAMSEVCKKNDLDSPKVLGFPSPKPIGVFIKRSIIIFR
SEQ ID NO: 12 GDSSEIYAEVVKGLGETLVGAYPGRALSFICKKRDLSNPQVLGNPSKPVGLFIRRSIIIFR
1381 1440
SEQ ID NO: 15 (gi3287270) SDSNGEDLEGYAGAGLYDSVPMDEEEKVVIDYSSDPLITDGNFRQTILSNIRAGHAIEE
SEQ ID NO: 10 SDSNGEDLEGYAGARLYDSVPMDEEDEVIVDYNNGLITDQGFQKSNLPSIAPAGHAIEE
SEQ ID NO: 12 SDSNGEDLEGNDGAGHYDSVPMGEPEKVVLDYSSDKLILDGSRQSISSIRAGNEIEE
1441 1466
SEQ ID NO: 15 (gi3287270) LYGSPQDIEGVVRDGGKIYVVQTRPQM
SEQ ID NO: 10 LYGSPQDVEGAVKEGKLYVVQTRPQM
SEQ ID NO: 12 LYGTPODIEGVIKDGKVVVQTRPQM

FIG. 1 (continued)

SEQUENCE LISTING

<110> E. I. DU PONT DE NEMOURS AND COMPANY
 <120> STARCH R1 PHOSPHORYLATION PROTEIN HOMOLOGS
 <130> BB-1158
 <140>
 <141>
 <160> 15
 <170> Microsoft Office 97
 <210> 1
 <211> 490
 <212> DNA
 <213> Arabidopsis
 <220>
 <221> unsure
 <222> (418)
 <220>
 <221> unsure
 <222> (429)
 <220>
 <221> unsure
 <222> (456)
 <220>
 <221> unsure
 <222> (473)
 <220>
 <221> unsure
 <222> (485)
 <220>
 <221> unsure
 <222> (487)
 <400> 1
 gttgaagaaa agaattgtaga gccactttctt gaggggtttgc ttgaagctcg tcaagagcta 60
 aggccacttc tgctgaagtc ccatgaccgc ctcaaggatc tgttattctt ggacctcgct 120
 cttgattcta ctgtcagaac agcgattgaa agaggatatg agcaattgaa tgatgctgga 180
 cctgagaaaa tcatgtactt catcagccta gttcttgaaa atcttgccct ctcttcagat 240
 gacaatgaag accttatata ctgcttgaag ggatggcaat ttgccctcga catgtgcaag 300
 agcaaaaaag atcactgggc tctgtatgca aaatctgttc ttgacagaag cccgactagc 360
 actggcaagc aaagctgaag aggtaccttg aaattctgca accatcggct gaatatcntg 420
 gatctgtcnt ggagtcgatc agtccggctg ttaatntatt actggaagaa atnattcgag 480
 ctggnntntgc 490
 <210> 2
 <211> 161
 <212> PRT
 <213> Arabidopsis thaliana
 <220>
 <221> UNSURE
 <222> (117)

<220>
 <221> UNSURE
 <222> (127)

<220>
 <221> UNSURE
 <222> (139)

<220>
 <221> UNSURE
 <222> (143)

<220>
 <221> UNSURE
 <222> (152)

<220>
 <221> UNSURE
 <222> (157)

<400> 2
 Val Glu Glu Lys Asn Val Glu Pro Leu Leu Glu Gly Leu Leu Glu Ala
 1 5 10 15
 Arg Gln Glu Leu Arg Pro Leu Leu Leu Lys Ser His Asp Arg Leu Lys
 20 25 30
 Asp Leu Leu Phe Leu Asp Leu Ala Leu Asp Ser Thr Val Arg Thr Ala
 35 40 45
 Ile Glu Arg Gly Tyr Glu Gln Leu Asn Asp Ala Gly Pro Glu Lys Ile
 50 55 60
 Met Tyr Phe Ile Ser Leu Val Leu Glu Asn Leu Ala Leu Ser Ser Asp
 65 70 75 80
 Asp Asn Glu Asp Leu Ile Tyr Cys Leu Lys Gly Trp Gln Phe Ala Leu
 85 90 95
 Asp Met Cys Lys Ser Lys Lys Asp His Trp Ala Leu Tyr Ala Lys Ser
 100 105 110
 Val Leu Asp Arg Xaa Arg Leu Ala Leu Ala Ser Lys Ala Glu Xaa Tyr
 115 120 125
 Leu Glu Ile Leu Gln Pro Ser Ala Glu Tyr Xaa Gly Ser Val Xaa Glu
 130 135 140
 Ser Ile Ser Pro Ala Val Asn Xaa Leu Leu Glu Glu Xaa Ile Arg Ala
 145 150 155 160
 Gly

<210> 3
 <211> 503
 <212> DNA
 <213> *Typha latifolia*

<220>
 <221> unsure
 <222> (359)

<220>
 <221> unsure
 <222> (380)

<220>
 <221> unsure
 <222> (389)

<220>
 <221> unsure
 <222> (409)

<220>
 <221> unsure
 <222> (429)

<220>
 <221> unsure
 <222> (459)

<220>
 <221> unsure
 <222> (479)

<220>
 <221> unsure
 <222> (481)

<220>
 <221> unsure
 <222> (493)

<400> 3
 agaaaaacag ttcttcaatt agcacctcca aatccgttgg tagaagagtt gaaggaaaaa 60
 atgcatgggtg ctggaatgcc atggcctggt gatgaagggtg aatctcgggtg ggaacaagca 120
 tggatggcaa taaaaaagggt atgggcttca aaatggaatg agagagcata cttcagcacc 180
 cgtaaaagtaa agttggatca tgactatctt tgcattggctg tcctgggtcca agaaattata 240
 aatgcaagat tatgcatttg tgatccatac tactaaccce tcaaccggag acgcatcaag 300
 agatatatgc tgagggtggtg aaaggactgg gagaagacac tagtgggaag cctacccang 360
 gtcgtgcctt aaagcttcan ctgttaagna agaaacgac ctaaactcnc caaaaagggtc 420
 ctgggtttnc ccaaaattaa acccaaattg gcctgttcna taaagaaaga tcaatccanc 480
 ntcaaaatta agnttcctaa tgg 503

<210> 4
 <211> 115
 <212> PRT
 <213> *Typha latifolia*

<220>
 <221> UNSURE
 <222> (83)

<220>
 <221> UNSURE
 <222> (100)

<400> 4
 Arg Lys Thr Val Leu Gln Leu Ala Pro Pro Asn Pro Leu Val Glu Glu
 1 5 10 15
 Leu Lys Glu Lys Met His Gly Ala Gly Met Pro Trp Pro Gly Asp Glu
 20 25 30
 Gly Glu Ser Arg Trp Glu Gln Ala Trp Met Ala Ile Lys Lys Val Trp
 35 40 45
 Ala Ser Lys Trp Asn Glu Arg Ala Tyr Phe Ser Thr Arg Lys Val Lys
 50 55 60

Leu Asp His Asp Tyr Leu Cys Met Ala Val Leu Val Gln Glu Ile Ile
65 70 75 80

Asn Ala Xaa Tyr Ala Phe Val Ile His Thr Thr Asn Pro Ser Thr Gly
85 90 95

Asp Ala Ser Xaa Ile Tyr Ala Glu Val Val Lys Gly Leu Gly Glu Asp
100 105 110

Thr Ser Gly
115

<210> 5
<211> 490
<212> DNA
<213> Curcuma zedoaria

<220>
<221> unsure
<222> (466)

<400> 5
aggtgatgtt ggtcagcgta tccgagatga aatattagtt ttacagagaa acaatgactg 60
caagggagga atgatggagg aatggcatca gaagctacat aacaacacta gcccagatga 120
tgttgtgata tgccaggcac ttattgatta tgtaaaagt gattttgaca tcagtgtgta 180
ctgggacagt ttgaataaaa atggaataac caaggaaagt ttgttgagct atgatcgtgc 240
tattcattct gaaccaagt tcaggagaga tcagaaagaa ggtcttttac gtgatctagg 300
aaactacatg aggacgttga aggcagttca ctctggtgca agatctcgag tctgccattg 360
ctacgtgtat gggttacaaa tctgagcgtc aagggtttta tggttggcgt tcaaataaac 420
cccgataggg ggattgccaa ctgggattcc ctgatctaaa ggaaantcaa tccaaaacat 480
gttgaaagat 490

<210> 6
<211> 144
<212> PRT
<213> Curcuma zedoaria

<220>
<221> UNSURE
<222> (114)

<220>
<221> UNSURE
<222> (130)

<400> 6
Gly Asp Val Gly Gln Arg Ile Arg Asp Glu Ile Leu Val Leu Gln Arg
1 5 10 15

Asn Asn Asp Cys Lys Gly Gly Met Met Glu Glu Trp His Gln Lys Leu
20 25 30

His Asn Asn Thr Ser Pro Asp Asp Val Val Ile Cys Gln Ala Leu Ile
35 40 45

Asp Tyr Val Lys Ser Asp Phe Asp Ile Ser Val Tyr Trp Asp Ser Leu
50 55 60

Asn Lys Asn Gly Ile Thr Lys Glu Arg Leu Leu Ser Tyr Asp Arg Ala
65 70 75 80

Ile His Ser Glu Pro Ser Phe Arg Arg Asp Gln Lys Glu Gly Leu Leu
85 90 95

Arg Asp Leu Gly Asn Tyr Met Arg Thr Leu Lys Ala Val His Ser Gly
100 105 110

Ala Xaa Leu Glu Ser Ala Ile Ala Thr Cys Met Gly Tyr Lys Ser Glu
 115 120 125

Arg Xaa Gly Phe Met Val Gly Val Gln Ile Asn Pro Asp Arg Gly Ile
 130 135 140

<210> 7
 <211> 487
 <212> DNA
 <213> Brachythecium oxycladon

<220>
 <221> unsure
 <222> (336)

<220>
 <221> unsure
 <222> (408)

<220>
 <221> unsure
 <222> (417)

<220>
 <221> unsure
 <222> (472)

<400> 7
 cacaccaaga tctggttggt gccaggtctc gtaatatagc caacctgcga ggcaaacttc 60
 cctcatggat tcattctcca acttcagcag cattgccatt tggagttttc gagaagggtt 120
 tagcagagcg catcaataag gatgtggcca cagaggttgc tgccctcagc aagcaacttg 180
 ctaatggtga ttttagtaag ctccaggatg ctggtgcaac ggtcttggga ctgaaagcac 240
 ctccagcgtt ggttgatgaa ttgaagaaaa ctttgaaaga ctgaggtatg ccgtggcctg 300
 gggatgaaag cgaggagaga tggactcaag cctgggctgc aatcaaaaagg gtgtgggcct 360
 caaaatggaa tgaaagagcc tacttcagta ctgcgaaaagc caagatanat cacaagngac 420
 ttgtgcatgg caagttatta gttcaagaga tcattcaagg ctgactatgc gntcgtcatt 480
 catacca 487

<210> 8
 <211> 159
 <212> PRT
 <213> Brachythecium oxycladon

<220>
 <221> UNSURE
 <222> (110)

<220>
 <221> UNSURE
 <222> (134)

<220>
 <221> UNSURE
 <222> (137)

<220>
 <221> UNSURE
 <222> (149)..(150)

<220>
 <221> UNSURE
 <222> (155)

<400> 8
 Asp Leu Val Gly Ala Lys Ser Arg Asn Ile Ala Asn Leu Arg Gly Lys
 1 5 10 15
 Leu Pro Ser Trp Ile His Leu Pro Thr Ser Ala Ala Leu Pro Phe Gly
 20 25 30
 Val Phe Glu Lys Val Leu Ala Glu Arg Ile Asn Lys Asp Val Ala Thr
 35 40 45
 Glu Val Ala Ala Leu Ser Lys Gln Leu Ala Asn Gly Asp Phe Ser Lys
 50 55 60
 Leu Gln Asp Ala Arg Ala Thr Val Leu Gly Leu Lys Ala Pro Pro Ala
 65 70 75 80
 Leu Val Asp Glu Leu Lys Lys Thr Leu Lys Asp Ser Gly Met Pro Trp
 85 90 95
 Pro Gly Asp Glu Ser Glu Glu Arg Trp Thr Gln Ala Trp Xaa Ala Ile
 100 105 110
 Lys Arg Val Trp Ala Ser Lys Trp Asn Glu Arg Ala Tyr Phe Ser Thr
 115 120 125
 Arg Lys Ala Lys Ile Xaa His Lys Xaa Leu Val His Gly Lys Leu Leu
 130 135 140
 Val Gln Glu Ile Xaa Xaa Ala Asp Tyr Ala Xaa Val Ile His Thr
 145 150 155

<210> 9
 <211> 1633
 <212> DNA
 <213> Oryza sativa

<220>
 <221> unsure
 <222> (874)

<220>
 <221> unsure
 <222> (876)

<400> 9
 tgggtacctgc taccctgtct gctcttctga atcggattga tccgtgttctt aggaatgttg 60
 cacagcttgg aagttggcag gttataagcc cagttgaagt atcaggttac attgtagtgg 120
 ttgatgaatt gcttgcgtgt caaaacaaat cctatgataa accaactatc cttgtggcaa 180
 agagtgtcaa ggagaggaa gaaataccag atggagtgtg tgggtgttatt acacctgata 240
 tgccagatgt tctctcccat gtatcagtc gagcaaggaa ttgcaagggt ttatttgcaa 300
 catgctttga tcctaaccac ttgtctgaac tccaaggaca tgatgggaaa gtgttttctc 360
 tcaaacctac ttctgcagat atcacctata gggagattcc agagagtga ctgcaatcag 420
 gttctctaaa tgcagaagct ggccaggcag tgccatctgt gtcattagtc aagaagaagt 480
 ttcttggaat atatgcaata tcagcagaag aattctctga ggaaatggtt ggggccaagt 540
 ctcgcaacgt agcatacctc aaaggaaaag taccctcatg gggtggtgtc cctacatcag 600
 ttgctgattcc atttgggacc ttgagaagg tttgtctga tgaaatcaat aagggaagtcg 660
 cgcaaacccat acaaatgctg aagggaaaac ttgctcaaga tgattttagt gctctaggcg 720
 aaatacggaa aactgttctc aatttaactg ctctactca actgatcaag gaactgaagg 780
 agaagatgct aggcctctga atgcctggc ctggagatga aggtgaccaa cggtgggagc 840
 aagcatggat ggcaattaaa aaggtttggg cgtnanaatg gaatgaaaga gcatatttta 900
 gcaactcgtaa ggtgaagctt gatcatgact acctttccat ggctgtactt gtacaagaaa 960
 ttgtcaatgc agactatgcc tttgtcattc atactactaa cccatcatcg ggagattcgt 1020
 ctgagatata tgctgaagtg gtgaaagggc ttggagaaac acttgtagga gcctatcctg 1080
 gtgcgcgccat gacgtttgta tgtaagaaaa acgaccttga ctctcccaa gtactgggtt 1140
 tcccaagcaa gccaatgggt gtcttcataa agagatcaat catctttcgt tcggattcca 1200
 acggtgagga tttagaaggg tatgctggag caagactgta tgatagtgtc cctatggatg 1260

```

aggaagatga agtcatagtc gactacaaca acggaccctt cattacagat cagggattcc 1320
aaaaatccaa cctcccagc attgcaccgg ctgggtcatgc cattgaggag ctttatgggt 1380
ccccacagga tgttgagggt gcagtgaagg aagggaagct atacgtagta cagacaagac 1440
cacagatgta atctatatgt atattttata gccaaagtcaa tcaggcaatg ttgtagagta 1500
agatatacgg gccgtgggac atgtataaca cgttacgccc ttttttttat tatttgcttt 1560
catactcaca atacactaat ttatagggct tattttatcg ccaaaaaaaaa aaaaaaaaaa 1620
aaaaaaaaaa aaa                                     1633

```

<210> 10
 <211> 482
 <212> PRT
 <213> Oryza sativa

<220>
 <221> UNSURE
 <222> (291)..(292)

<400> 10
 Val Pro Ala Thr Leu Ser Ala Leu Leu Asn Arg Ile Asp Pro Val Leu
 1 5 10 15
 Arg Asn Val Ala Gln Leu Gly Ser Trp Gln Val Ile Ser Pro Val Glu
 20 25 30
 Val Ser Gly Tyr Ile Val Val Val Asp Glu Leu Leu Ala Val Gln Asn
 35 40 45
 Lys Ser Tyr Asp Lys Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly
 50 55 60
 Glu Glu Glu Ile Pro Asp Gly Val Val Gly Val Ile Thr Pro Asp Met
 65 70 75 80
 Pro Asp Val Leu Ser His Val Ser Val Arg Ala Arg Asn Cys Lys Val
 85 90 95
 Leu Phe Ala Thr Cys Phe Asp Pro Asn Thr Leu Ser Glu Leu Gln Gly
 100 105 110
 His Asp Gly Lys Val Phe Ser Phe Lys Pro Thr Ser Ala Asp Ile Thr
 115 120 125
 Tyr Arg Glu Ile Pro Glu Ser Glu Leu Gln Ser Gly Ser Leu Asn Ala
 130 135 140
 Glu Ala Gly Gln Ala Val Pro Ser Val Ser Leu Val Lys Lys Lys Phe
 145 150 155 160
 Leu Gly Lys Tyr Ala Ile Ser Ala Glu Glu Phe Ser Glu Glu Met Val
 165 170 175
 Gly Ala Lys Ser Arg Asn Val Ala Tyr Leu Lys Gly Lys Val Pro Ser
 180 185 190
 Trp Val Gly Val Pro Thr Ser Val Ala Ile Pro Phe Gly Thr Phe Glu
 195 200 205
 Lys Val Leu Ser Asp Glu Ile Asn Lys Glu Val Ala Gln Thr Ile Gln
 210 215 220
 Met Leu Lys Gly Lys Leu Ala Gln Asp Asp Phe Ser Ala Leu Gly Glu
 225 230 235 240
 Ile Arg Lys Thr Val Leu Asn Leu Thr Ala Pro Thr Gln Leu Ile Lys
 245 250 255

Glu Leu Lys Glu Lys Met Leu Gly Ser Gly Met Pro Trp Pro Gly Asp
 260 265 270
 Glu Gly Asp Gln Arg Trp Glu Gln Ala Trp Met Ala Ile Lys Lys Val
 275 280 285
 Trp Ala Xaa Xaa Trp Asn Glu Arg Ala Tyr Phe Ser Thr Arg Lys Val
 290 295 300
 Lys Leu Asp His Asp Tyr Leu Ser Met Ala Val Leu Val Gln Glu Ile
 305 310 315 320
 Val Asn Ala Asp Tyr Ala Phe Val Ile His Thr Thr Asn Pro Ser Ser
 325 330 335
 Gly Asp Ser Ser Glu Ile Tyr Ala Glu Val Val Lys Gly Leu Gly Glu
 340 345 350
 Thr Leu Val Gly Ala Tyr Pro Gly Arg Ala Met Ser Phe Val Cys Lys
 355 360 365
 Lys Asn Asp Leu Asp Ser Pro Lys Val Leu Gly Phe Pro Ser Lys Pro
 370 375 380
 Ile Gly Val Phe Ile Lys Arg Ser Ile Ile Phe Arg Ser Asp Ser Asn
 385 390 395 400
 Gly Glu Asp Leu Glu Gly Tyr Ala Gly Ala Arg Leu Tyr Asp Ser Val
 405 410 415
 Pro Met Asp Glu Glu Asp Glu Val Ile Val Asp Tyr Asn Asn Gly Pro
 420 425 430
 Leu Ile Thr Asp Gln Gly Phe Gln Lys Ser Asn Leu Pro Ser Ile Ala
 435 440 445
 Pro Ala Gly His Ala Ile Glu Glu Leu Tyr Gly Ser Pro Gln Asp Val
 450 455 460
 Glu Gly Ala Val Lys Glu Gly Lys Leu Tyr Val Val Gln Thr Arg Pro
 465 470 475 480

Gln Met

<210> 11
 <211> 517
 <212> DNA
 <213> Glycine max

<220>
 <221> unsure
 <222> (371)

<220>
 <221> unsure
 <222> (494)

<400> 11
 aaggtacagc caagttcttg ttgaataaaa tagcggaaat ggaaagttag gcacaaaagt 60
 ccttcacgca tcgatttaac attgcatcgg atttgataga tgaagctaaa aatgctggtc 120
 aacaaggtct tgcggggatt ttggtgtgga tgagattcat ggctactagg cagctcatat 180
 ggaacaaaaa ttacaatgtg aagccacgtg agataagtaa agcacaggat aggttacag 240
 acttgctcca ggatgtttat gcaagttacc cacagtatag ggaaattgtg aggatgatct 300
 tgctcgactgt tggctcgtgga ggtgaaggag atgtcggaca gaggattcgg gatgaaatcc 360
 ttgttatcca ngagaaataa tgattgtaaa ggtggaatga tggaggaatg gcaccagaaa 420

ttacacaata atactagtcc tgatgatgtt gtaatctgtc aagcactaat tgattatata 480
 aatagtgact ttgntattgg tggttactgg caaacat 517

<210> 12
 <211> 171
 <212> PRT
 <213> Glycine max

<220>
 <221> UNSURE
 <222> (123)

<220>
 <221> UNSURE
 <222> (164)

<400> 12
 Gly Thr Ala Lys Phe Leu Leu Asn Lys Ile Ala Glu Met Glu Ser Glu
 1 5 10 15
 Ala Gln Lys Ser Phe Met His Arg Phe Asn Ile Ala Ser Asp Leu Ile
 20 25 30
 Asp Glu Ala Lys Asn Ala Gly Gln Gln Gly Leu Ala Gly Ile Leu Val
 35 40 45
 Trp Met Arg Phe Met Ala Thr Arg Gln Leu Ile Trp Asn Lys Asn Tyr
 50 55 60
 Asn Val Lys Pro Arg Glu Ile Ser Lys Ala Gln Asp Arg Leu Thr Asp
 65 70 75 80
 Leu Leu Gln Asp Val Tyr Ala Ser Tyr Pro Gln Tyr Arg Glu Ile Val
 85 90 95
 Arg Met Ile Leu Ser Thr Val Gly Arg Gly Gly Glu Gly Asp Val Gly
 100 105 110
 Gln Arg Ile Arg Asp Glu Ile Leu Val Ile Xaa Arg Asn Asn Asp Cys
 115 120 125
 Lys Gly Gly Met Met Glu Glu Trp His Gln Lys Leu His Asn Asn Thr
 130 135 140
 Ser Pro Asp Asp Val Val Ile Cys Gln Ala Leu Ile Asp Tyr Ile Asn
 145 150 155 160
 Ser Asp Phe Xaa Ile Gly Val Tyr Trp Gln Thr
 165 170

<210> 13
 <211> 2080
 <212> DNA
 <213> Glycine max

<400> 13
 aaataatgta cttcattagc ttgggttcttg aaaatctcgc actttcatcg gatgacaatg 60
 aagatcttat ctactgtttg aagggatggg atgttgccctt aagcatgtgc aagattaaaag 120
 atactcattg ggcattgtac gcaaaaatcag tccttgacag aaccctgctt gcactaaciaa 180
 acaaggctca ttatataccag gaaattctgc aaccatcggc agaatatctt ggatcactgc 240
 ttggcgtgga caaatgggcc gtggaaatat ttactgaaga aattatccgt gctggatctg 300
 ctgcttcttt gtctactctt ctaaatacgac tggatcctgt gctccgaaag acagctcatc 360
 ttggaagctg gcagggttatt agcccagttg aaactgttg atatgttgag gtcatagatg 420
 agttgcttgc tgttcaaaac aaatcatatg agcgacctac aattttgata gccaaagatg 480
 tgagaggaga ggaagaaatt ccagatggta cagttgctgt cctgacacct gatatgcccg 540
 atgtcctatc ccatgtatct gtacgagcaa gaaatagcaa ggtgtgtttt gctacatgct 600

```

ttgatcccaa tctcctggct aacctccaag aaaataaagg aaagcttttg cgcttaaagc 660
caacatctgc tgatgtagtt tatagttagg tcaagggaagg tgagttaatt gatgacaaat 720
caactcaact caaagatggt gggtctgtgt caccatatac tctggcccga aagaagttta 780
gtggttagata tgctgtctca tctgaagaat tcaactggtga aatggttgga gctaaatctc 840
gtaatatctc ttatttaaaa gggaaagtag cttcttggtg tggaattcct acctcggttg 900
ccataccatt tggagttttc gaacatgttc tttctgataa accaaaccag gcagtggctg 960
agagggtcaa taatttgaaa aagaagttaa ttgagggaga cttcagtggt ctcaaggaga 1020
ttcgtgaaac agttctacaa ttgaatgcac catcccatth ggtagaggag ttgaaaacta 1080
aaatgaagag ttctggaatg ccgtggccgg gtgatgaagg tgaacaacga tgggagcaag 1140
cttgatagc tataaaaaag gtgtggggct ctaagtggaa tgaaagagca tacttcagca 1200
caagaaaagt gaaactcgac cacgaatata tttccatggc agtccctgtt caagaagtga 1260
taaattgctga ctatgctttt gtcattccaca caactaaccg tgccctctga gattcatcgg 1320
aaatatatgc tgaggtggta aagggacttg gagaaacact ggttgaggct tatccaggtc 1380
gtgcttttag ttttatctgc aagaaacgtg atttgaactc tcctcaggtc ttaggtaatc 1440
ctagcaaacc tgcggccta tttataagac ggtcaattat ttttcgatct gattccaatg 1500
gtgaagatct agaaggtaat gtggtgcag gtcattatga cagtgtccca atgggtgaac 1560
ccgagaaggt ggtgcttgat tattcttcag acaaaactgat ccttgatggc agttttcgcc 1620
agtcaatctt gtccagcatt gcccggtgcag gaaatgaaat tgaagagttg tatggcactc 1680
ctcaggacat tgaaggtgtc atcaaggatg gaaaagtcta tgtgtccag accagaccac 1740
aaatgtagac ctccatacct atgtctttta agccaactac ctcaactatg ttctatgttc 1800
attcccggtc aacatggcgt ttcaaactgt gccgtggcag cttctgagag tttaagagta 1860
accgcgggga ttaccaaatt tggccttata gatttattac acgtgatata ttgaaaatta 1920
aggaataatt tataagtgtg taaacatgga ataattgtaa ttaattaaaa aattagatgg 1980
tcttattctt tttccctact atatatattg tatgtactta cttcttccta attaaaattg 2040
ctattcaaaag taaaaaaaaa aaagggggcg ccggtaccca 2080

```

<210> 14
 <211> 581
 <212> PRT
 <213> Glycine max

<400> 14
 Ile Met Tyr Phe Ile Ser Leu Val Leu Glu Asn Leu Ala Leu Ser Ser
 1 5 10 15
 Asp Asp Asn Glu Asp Leu Ile Tyr Cys Leu Lys Gly Trp Asp Val Ala
 20 25 30
 Leu Ser Met Cys Lys Ile Lys Asp Thr His Trp Ala Leu Tyr Ala Lys
 35 40 45
 Ser Val Leu Asp Arg Thr Arg Leu Ala Leu Thr Asn Lys Ala His Leu
 50 55 60
 Tyr Gln Glu Ile Leu Gln Pro Ser Ala Glu Tyr Leu Gly Ser Leu Leu
 65 70 75 80
 Gly Val Asp Lys Trp Ala Val Glu Ile Phe Thr Glu Glu Ile Ile Arg
 85 90 95
 Ala Gly Ser Ala Ala Ser Leu Ser Thr Leu Leu Asn Arg Leu Asp Pro
 100 105 110
 Val Leu Arg Lys Thr Ala His Leu Gly Ser Trp Gln Val Ile Ser Pro
 115 120 125
 Val Glu Thr Val Gly Tyr Val Glu Val Ile Asp Glu Leu Leu Ala Val
 130 135 140
 Gln Asn Lys Ser Tyr Glu Arg Pro Thr Ile Leu Ile Ala Lys Ser Val
 145 150 155 160
 Arg Gly Glu Glu Glu Ile Pro Asp Gly Thr Val Ala Val Leu Thr Pro
 165 170 175

Asp Met Pro Asp Val Leu Ser His Val Ser Val Arg Ala Arg Asn Ser
 180 185 190
 Lys Val Cys Phe Ala Thr Cys Phe Asp Pro Asn Ile Leu Ala Asn Leu
 195 200 205
 Gln Glu Asn Lys Gly Lys Leu Leu Arg Leu Lys Pro Thr Ser Ala Asp
 210 215 220
 Val Val Tyr Ser Glu Val Lys Glu Gly Glu Leu Ile Asp Asp Lys Ser
 225 230 235 240
 Thr Gln Leu Lys Asp Val Gly Ser Val Ser Pro Ile Ser Leu Ala Arg
 245 250 255
 Lys Lys Phe Ser Gly Arg Tyr Ala Val Ser Ser Glu Glu Phe Thr Gly
 260 265 270
 Glu Met Val Gly Ala Lys Ser Arg Asn Ile Ser Tyr Leu Lys Gly Lys
 275 280 285
 Val Ala Ser Trp Ile Gly Ile Pro Thr Ser Val Ala Ile Pro Phe Gly
 290 295 300
 Val Phe Glu His Val Leu Ser Asp Lys Pro Asn Gln Ala Val Ala Glu
 305 310 315 320
 Arg Val Asn Asn Leu Lys Lys Lys Leu Ile Glu Gly Asp Phe Ser Val
 325 330 335
 Leu Lys Glu Ile Arg Glu Thr Val Leu Gln Leu Asn Ala Pro Ser His
 340 345 350
 Leu Val Glu Glu Leu Lys Thr Lys Met Lys Ser Ser Gly Met Pro Trp
 355 360 365
 Pro Gly Asp Glu Gly Glu Gln Arg Trp Glu Gln Ala Trp Ile Ala Ile
 370 375 380
 Lys Lys Val Trp Gly Ser Lys Trp Asn Glu Arg Ala Tyr Phe Ser Thr
 385 390 395 400
 Arg Lys Val Lys Leu Asp His Glu Tyr Leu Ser Met Ala Val Leu Val
 405 410 415
 Gln Glu Val Ile Asn Ala Asp Tyr Ala Phe Val Ile His Thr Thr Asn
 420 425 430
 Pro Ala Ser Gly Asp Ser Ser Glu Ile Tyr Ala Glu Val Val Lys Gly
 435 440 445
 Leu Gly Glu Thr Leu Val Gly Ala Tyr Pro Gly Arg Ala Leu Ser Phe
 450 455 460
 Ile Cys Lys Lys Arg Asp Leu Asn Ser Pro Gln Val Leu Gly Asn Pro
 465 470 475 480
 Ser Lys Pro Val Gly Leu Phe Ile Arg Arg Ser Ile Ile Phe Arg Ser
 485 490 495
 Asp Ser Asn Gly Glu Asp Leu Glu Gly Asn Asp Gly Ala Gly His Tyr
 500 505 510
 Asp Ser Val Pro Met Gly Glu Pro Glu Lys Val Val Leu Asp Tyr Ser
 515 520 525

Ser Asp Lys Leu Ile Leu Asp Gly Ser Phe Arg Gln Ser Ile Leu Ser
530 535 540

Ser Ile Ala Arg Ala Gly Asn Glu Ile Glu Glu Leu Tyr Gly Thr Pro
545 550 555 560

Gln Asp Ile Glu Gly Val Ile Lys Asp Gly Lys Val Tyr Val Val Gln
565 570 575

Thr Arg Pro Gln Met
580

<210> 15

<211> 1464

<212> PRT

<213> Solanum tuberosum

<400> 15

Met Ser Asn Ser Leu Gly Asn Asn Leu Leu Tyr Gln Gly Phe Leu Thr
1 5 10 15

Ser Thr Val Leu Glu His Lys Ser Arg Ile Ser Pro Pro Cys Val Gly
20 25 30

Gly Asn Ser Leu Phe Gln Gln Gln Val Ile Ser Lys Ser Pro Leu Ser
35 40 45

Thr Glu Phe Arg Gly Asn Arg Leu Lys Val Gln Lys Lys Lys Ile Pro
50 55 60

Met Glu Lys Lys Arg Ala Phe Ser Ser Ser Pro His Ala Val Leu Thr
65 70 75 80

Thr Asp Thr Ser Ser Glu Leu Ala Glu Lys Phe Ser Leu Gly Gly Asn
85 90 95

Ile Glu Leu Gln Val Asp Val Arg Pro Pro Thr Ser Gly Asp Val Ser
100 105 110

Phe Val Asp Phe Gln Val Thr Asn Gly Ser Asp Lys Leu Phe Leu His
115 120 125

Trp Gly Ala Val Lys Phe Gly Lys Glu Thr Trp Ser Leu Pro Asn Asp
130 135 140

Arg Pro Asp Gly Thr Lys Val Tyr Lys Asn Lys Ala Leu Arg Thr Pro
145 150 155 160

Phe Val Lys Ser Gly Ser Asn Ser Ile Leu Arg Leu Glu Ile Arg Asp
165 170 175

Thr Ala Ile Glu Ala Ile Glu Phe Leu Ile Tyr Asp Glu Ala His Asp
180 185 190

Lys Trp Ile Lys Asn Asn Gly Gly Asn Phe Arg Val Lys Leu Ser Arg
195 200 205

Lys Glu Ile Arg Gly Pro Asp Val Ser Val Pro Glu Glu Leu Val Gln
210 215 220

Ile Gln Ser Tyr Leu Arg Trp Glu Arg Lys Gly Lys Gln Asn Tyr Pro
225 230 235 240

Pro Glu Lys Glu Lys Glu Glu Tyr Glu Ala Ala Arg Thr Val Leu Gln
 245 250 255
 Glu Glu Ile Ala Arg Gly Ala Ser Ile Gln Asp Ile Arg Ala Arg Leu
 260 265 270
 Thr Lys Thr Asn Asp Lys Ser Gln Ser Lys Glu Glu Pro Leu His Val
 275 280 285
 Thr Lys Ser Asp Ile Pro Asp Asp Leu Ala Gln Ala Gln Ala Tyr Ile
 290 295 300
 Arg Trp Glu Lys Ala Gly Lys Pro Asn Tyr Pro Pro Glu Lys Gln Ile
 305 310 315 320
 Glu Glu Leu Glu Glu Ala Arg Arg Glu Leu Gln Leu Glu Leu Glu Lys
 325 330 335
 Gly Ile Thr Leu Asp Glu Leu Arg Lys Thr Ile Thr Lys Gly Glu Ile
 340 345 350
 Lys Thr Lys Val Glu Lys His Leu Lys Arg Ser Ser Phe Ala Val Glu
 355 360 365
 Arg Ile Gln Arg Lys Lys Arg Asp Phe Gly His Leu Ile Asn Lys Tyr
 370 375 380
 Thr Ser Ser Pro Ala Val Gln Val Gln Lys Val Leu Glu Glu Pro Pro
 385 390 395 400
 Ala Leu Ser Lys Ile Lys Leu Tyr Ala Lys Glu Lys Glu Glu Gln Ile
 405 410 415
 Asp Asp Pro Ile Leu Asn Lys Lys Ile Phe Lys Val Asp Asp Gly Glu
 420 425 430
 Leu Leu Val Leu Val Ala Lys Ser Ser Gly Lys Thr Lys Val His Leu
 435 440 445
 Ala Thr Asp Leu Asn Gln Pro Ile Thr Leu His Trp Ala Leu Ser Lys
 450 455 460
 Ser Pro Gly Glu Trp Met Val Pro Pro Ser Ser Ile Leu Pro Pro Gly
 465 470 475 480
 Ser Ile Ile Leu Asp Lys Ala Ala Glu Thr Pro Phe Ser Ala Ser Ser
 485 490 495
 Ser Asp Gly Leu Thr Ser Lys Val Gln Ser Leu Asp Ile Val Ile Glu
 500 505 510
 Asp Gly Asn Phe Val Gly Met Pro Phe Val Leu Leu Ser Gly Glu Lys
 515 520 525
 Trp Ile Lys Asn Gln Gly Ser Asp Phe Tyr Val Gly Phe Ser Ala Ala
 530 535 540
 Ser Lys Leu Ala Leu Lys Ala Ala Gly Asp Gly Ser Gly Thr Ala Lys
 545 550 555 560
 Ser Leu Leu Asp Lys Ile Ala Asp Met Glu Ser Glu Ala Gln Lys Ser
 565 570 575
 Phe Met His Arg Phe Asn Ile Ala Ala Asp Leu Ile Glu Asp Ala Thr
 580 585 590

Ser Ala Gly Glu Leu Gly Phe Ala Gly Ile Leu Val Trp Met Arg Phe
 595 600 605
 Met Ala Thr Arg Gln Leu Ile Trp Asn Lys Asn Tyr Asn Val Lys Pro
 610 615 620
 Arg Glu Ile Ser Lys Ala Gln Asp Arg Leu Thr Asp Leu Leu Gln Asn
 625 630 635 640
 Ala Phe Thr Ser His Pro Gln Tyr Arg Glu Ile Leu Arg Met Ile Met
 645 650 655
 Ser Thr Val Gly Arg Gly Gly Glu Gly Asp Val Gly Gln Arg Ile Arg
 660 665 670
 Asp Glu Ile Leu Val Ile Gln Arg Asn Asn Asp Cys Lys Gly Gly Met
 675 680 685
 Met Gln Glu Trp His Gln Lys Leu His Asn Asn Thr Ser Pro Asp Asp
 690 695 700
 Val Val Ile Cys Gln Ala Leu Ile Asp Tyr Ile Lys Ser Asp Phe Asp
 705 710 715 720
 Leu Gly Val Tyr Trp Lys Thr Leu Asn Glu Asn Gly Ile Thr Lys Glu
 725 730 735
 Arg Leu Leu Ser Tyr Asp Arg Ala Ile His Ser Glu Pro Asn Phe Arg
 740 745 750
 Gly Asp Gln Lys Gly Gly Leu Leu Arg Asp Leu Gly His Tyr Met Arg
 755 760 765
 Thr Leu Lys Ala Val His Ser Gly Ala Asp Leu Glu Ser Ala Ile Ala
 770 775 780
 Asn Cys Met Gly Tyr Lys Thr Glu Gly Glu Gly Phe Met Val Gly Val
 785 790 795 800
 Gln Ile Asn Pro Val Ser Gly Leu Pro Ser Gly Phe Gln Asp Leu Leu
 805 810 815
 His Phe Val Leu Asp His Val Glu Asp Lys Asn Val Glu Thr Leu Leu
 820 825 830
 Glu Arg Leu Leu Glu Ala Arg Glu Glu Leu Arg Pro Leu Leu Leu Lys
 835 840 845
 Pro Asn Asn Arg Leu Lys Asp Leu Leu Phe Leu Asp Ile Ala Leu Asp
 850 855 860
 Ser Thr Val Arg Thr Ala Val Glu Arg Gly Tyr Glu Glu Leu Asn Asn
 865 870 875 880
 Ala Asn Pro Glu Lys Ile Met Tyr Phe Ile Ser Leu Val Leu Glu Asn
 885 890 895
 Leu Ala Leu Ser Val Asp Asp Asn Glu Asp Leu Val Tyr Cys Leu Lys
 900 905 910
 Gly Trp Asn Gln Ala Leu Ser Met Ser Asn Gly Gly Asp Asn His Trp
 915 920 925
 Ala Leu Phe Ala Lys Ala Val Leu Asp Arg Thr Arg Leu Ala Leu Ala
 930 935 940

Ser Lys Ala Glu Trp Tyr His His Leu Leu Gln Pro Ser Ala Glu Tyr
 945 950 955 960
 Leu Gly Ser Ile Leu Gly Val Asp Gln Trp Ala Leu Asn Ile Phe Thr
 965 970 975
 Glu Glu Ile Ile Arg Ala Gly Ser Ala Ala Ser Leu Ser Ser Leu Leu
 980 985 990
 Asn Arg Leu Asp Pro Val Leu Arg Lys Thr Ala Asn Leu Gly Ser Trp
 995 1000 1005
 Gln Ile Ile Ser Pro Val Glu Ala Val Gly Tyr Val Val Val Val Asp
 1010 1015 1020
 Glu Leu Leu Ser Val Gln Asn Glu Ile Tyr Glu Lys Pro Thr Ile Leu
 1025 1030 1035 1040
 Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro Asp Gly Ala Val
 1045 1050 1055
 Ala Leu Ile Thr Pro Asp Met Pro Asp Val Leu Ser His Val Ser Val
 1060 1065 1070
 Arg Ala Arg Asn Gly Lys Val Cys Phe Ala Thr Cys Phe Asp Pro Asn
 1075 1080 1085
 Ile Leu Ala Asp Leu Gln Ala Lys Glu Gly Arg Ile Leu Leu Leu Lys
 1090 1095 1100
 Pro Thr Pro Ser Asp Ile Ile Tyr Ser Glu Val Asn Glu Ile Glu Leu
 1105 1110 1115 1120
 Gln Ser Ser Ser Asn Leu Val Glu Ala Glu Thr Ser Ala Thr Leu Arg
 1125 1130 1135
 Leu Val Lys Lys Gln Phe Gly Gly Cys Tyr Ala Ile Ser Ala Asp Glu
 1140 1145 1150
 Phe Thr Ser Glu Met Val Gly Ala Lys Ser Arg Asn Ile Ala Tyr Leu
 1155 1160 1165
 Lys Gly Lys Val Pro Ser Ser Val Gly Ile Pro Thr Ser Val Ala Leu
 1170 1175 1180
 Pro Phe Gly Val Phe Glu Lys Val Leu Ser Asp Asp Ile Asn Gln Gly
 1185 1190 1195 1200
 Val Ala Lys Glu Leu Gln Ile Leu Met Lys Lys Leu Ser Glu Gly Asp
 1205 1210 1215
 Phe Ser Ala Leu Gly Glu Ile Arg Thr Thr Val Leu Asp Leu Ser Ala
 1220 1225 1230
 Pro Ala Gln Leu Val Lys Glu Leu Lys Glu Lys Met Gln Gly Ser Gly
 1235 1240 1245
 Met Pro Trp Pro Gly Asp Glu Gly Pro Lys Arg Trp Glu Gln Ala Trp
 1250 1255 1260
 Met Ala Ile Lys Lys Val Trp Ala Ser Lys Trp Asn Glu Arg Ala Tyr
 1265 1270 1275 1280
 Phe Ser Thr Arg Lys Val Lys Leu Asp His Asp Tyr Leu Cys Met Ala
 1285 1290 1295

Val Leu Val Gln Glu Ile Ile Asn Ala Asp Tyr Ala Phe Val Ile His
1300 1305 1310

Thr Thr Asn Pro Ser Ser Gly Asp Asp Ser Glu Ile Tyr Ala Glu Val
1315 1320 1325

Val Arg Gly Leu Gly Glu Thr Leu Val Gly Ala Tyr Pro Gly Arg Ala
1330 1335 1340

Leu Ser Phe Ile Cys Lys Lys Lys Asp Leu Asn Ser Pro Gln Val Leu
1345 1350 1355 1360

Gly Tyr Pro Ser Lys Pro Ile Gly Leu Phe Ile Lys Arg Ser Ile Ile
1365 1370 1375

Phe Arg Ser Asp Ser Asn Gly Glu Asp Leu Glu Gly Tyr Ala Gly Ala
1380 1385 1390

Gly Leu Tyr Asp Ser Val Pro Met Asp Glu Glu Glu Lys Val Val Ile
1395 1400 1405

Asp Tyr Ser Ser Asp Pro Leu Ile Thr Asp Gly Asn Phe Arg Gln Thr
1410 1415 1420

Ile Leu Ser Asn Ile Ala Arg Ala Gly His Ala Ile Glu Glu Leu Tyr
1425 1430 1435 1440

Gly Ser Pro Gln Asp Ile Glu Gly Val Val Arg Asp Gly Lys Ile Tyr
1445 1450 1455

Val Val Gln Thr Arg Pro Gln Met
1460

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 99/07639

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/52 C12N9/00 C12N15/82 C12N5/10 //A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 11188 A (KOSSMANN JENS ; LORBERTH RUTH (DE); PLANTTEC BIOTECHNOLOGIE GMBH (D) 27 March 1997 (1997-03-27) the whole document	1-7,9
X	SAKAKI, T., ET AL. : "rice cDNA from panicle at flowering stage" EMBL SEQUENCE DATA LIBRARY, 19 September 1997 (1997-09-19), XP002111458 heidelberg, germany accession no. C71741	1,2
P, X	WO 98 27212 A (EMMERMANN MICHAEL ; KOSSMANN JENS (DE); PLANTTEC BIOTECHNOLOGIE GMBH) 25 June 1998 (1998-06-25) the whole document	1-7,9
-/--		

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

- 6 August 1999

Date of mailing of the international search report

19/08/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Holtorf, S

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/07639

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 98 53085 A (SEYMOUR GRAHAM BARRON ;ZENECA LTD (GB); BIRD COLIN ROGER (GB); MED) 26 November 1998 (1998-11-26) the whole document	1-6
P,X	--- LORBERTH,R., ET AL. : "inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening" NATURE BIOTECHNOLOGY, vol. 16, May 1998 (1998-05), pages 473-477, XP002111459. cited in the application the whole document	1-3,6
P,X	--- ALCALA,J., ET AL. : "generation of ESTs from tomato carpel tissue" EMBL SEQUENCE DATA LIBRARY, 17 March 1999 (1999-03-17), XP002111460 heidelberg, germany accession no. AI489255	1,2
P,X	--- LIN, X., ET AL. : "arabidopsis thaliana "IGF" BAC "F23B24" genomic sequence near marker "mi473"" EMBL SEQUENCE DATA LIBRARY, 27 October 1998 (1998-10-27), XP002111461 heidelberg, germany accession no. AC005861	1,2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/ 07639

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99 07639

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-10 partially

Nucleic acid fragments that encode starch R1 phosphorylase isolated from Arabidopsis and the corresponding deduced amino acid sequence consisting of SEQID 1 and 2, respectively.

2. Claims: 1-10 partially

Nucleic acid fragments that encode starch R1 phosphorylase isolated from ginger and the corresponding deduced amino acid sequence consisting of SEQID 3 and 4, respectively.

3. Claims: 1-10 partially

Nucleic acid fragments that encode starch R1 phosphorylase isolated from moss and the corresponding deduced amino acid sequence consisting of SEQID 5 and 6, respectively.

4. Claims: 1-10 partially

Nucleic acid fragments that encode starch R1 phosphorylase isolated from cattail and the corresponding deduced amino acid sequence consisting of SEQID 7 and 8, respectively.

5. Claims: 1-10 partially

Nucleic acid fragments that encode starch R1 phosphorylase isolated from rice and the corresponding deduced amino acid sequence consisting of SEQID 9 and 10, respectively.

6. Claims: 1-10 partially

Nucleic acid fragments that encode starch R1 phosphorylase isolated from soybean and the corresponding deduced amino acid sequence consisting of SEQID 11,13 and 12,14, respectively.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99 07639

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claim 5 does not state the exact range of SEQID-numbers for which coverage is sought.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Inter national Application No

PCT/US 99/07639

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9711188 A	27-03-1997	DE 19534759 A DE 19547733 A AU 7131396 A CA 2231774 A EP 0851934 A HU 9900510 A	20-03-1997 26-06-1997 09-04-1997 27-03-1997 08-07-1998 28-05-1999
WO 9827212 A	25-06-1998	DE 19653176 A AU 5857798 A	25-06-1998 15-07-1998
WO 9853085 A	26-11-1998	AU 7225798 A	11-12-1998